09/044,696-AC-1

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/00, 15/31, A61K 39/106 A61K 39/108, C12P 21/02 C12N 1/21 // (C12N 1/21 C12R 1:19)

A1

(11) International Publication Number: WO 93/13202

(43) International Publication Date: 8 July 1993 (08.07.93)

(21) International Application Number: PCT/EP92/03016

(22) International Filing Date: 30 December 1992 (30.12.92)

(30) Priority data:
MI91A03513 31 December 1991 (31.12.91) IT

(71) Applicant *(for all designated States except US)*: BIOCINE SCLAVO SPA [IT/IT]; Via Fiorentina, 1, I-53100 Siena (IT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : DOMENIGHINI, Mario [IT/IT]; Via Ungaretti, 17, I-53010 Quercegrossa (IT). RAPPUOLI, Rino [IT/IT]; Via Calamandrei, 39, I-53010 Quercegrossa (IT). PIZZA, Mariagrazia [IT/IT]; Via Colombini, 30, I-43100 Siena (IT). HOL, Wim [NL/US]; 18332 57th Avenue, N.E., Seattle, WA 98155 (US).

(74) Agent: HALLYBONE, Huw, George; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

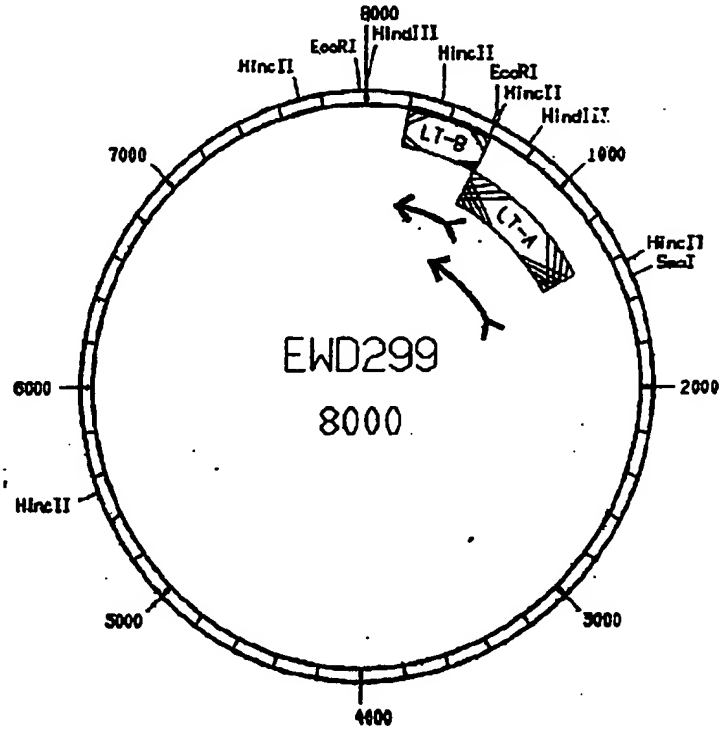
(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN LT, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES

(57) Abstract

An immunogenic detoxified protein comprising the amino acid sequence of subunit A of cholera toxin (CT-A) or subunit A of an *Escherichia coli* heat labile toxin (LT-A) or a fragment thereof wherein one or more amino acids at, or in positions corresponding to Val-53, Ser-63, Val-97, Tyr-104 or Pro-106 are replaced with another amino acid or deleted. Examples of specific replacements include Val-53-Asp, Val-53-Glu, Val-53-Tyr, Ser-63-Lys, Val-97-Lys, Val-97-Tyr, Tyr-104-Lys, Tyr-104-Asp, Tyr-104-Ser, Pro-106-Ser. The immunogenic detoxified protein is useful as vaccine for *Vibrio cholerae* or an enterotoxigenic strain of *Escherichia coli* and is produced by recombinant DNA means by site-directed mutagenesis.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	Côte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML	Mali	VN	Viet Nam
FI	Finland	MN	Mongolia		

Immunogenic detoxified mutants of cholera toxin and of the toxin LT, their preparation and their use for the preparation of vaccines

Field of the Invention

The present invention relates to immunogenic detoxified proteins of cholera toxins (CT), or of heat labile toxins (LT) produced by the enterotoxigenic strains of *Escherichia coli* (*E.coli*) having substitutions at one or more of amino acids Val-53, Ser-63, Val-97, Tyr-104 or Pro-106 and to their use in vaccines which are useful for the prevention or treatment of cholera or enterotoxigenic *E.Coli* infections. The proteins can be suitably produced using recombinant DNA techniques by site-directed mutagenesis of DNA encoding the wild type toxins.

Background to the Invention

Cholera is a contagious disease widely distributed in the world, in particular in the Third World, where, in certain areas, it is endemic. The serious disorders which develop in the intestinal system prove fatal in a high percentage of the recorded cases of the disease.

The etiological agent of cholera is the Gram-negative microorganism *Vibrio cholerae* (*V.cholerae*). This colonises the intestinal tract of individuals who have come into contact with it through ingestion of contaminated food or water, and multiplies to very high concentrations. The principal symptom is severe diarrhoea as a result of which the patient can lose as much as 10-15 litres of liquids per day via the faeces. As a result of the severe dehydration and loss of electrolytes, the patient does not withstand the infection in 50-60% of cases, and dies. The diarrhoea caused by *V.cholerae* is due to the secretion of cholera toxin, CT, which acts by stimulating the activity of the adenylate cyclase enzyme so as to induce disturbances at cell level.

Alth ugh cholera can be effectively cur d by controlled and intense rehydrati n, the distribution of a vaccine is desirable with a view to complete control and futur eradication of the disease.

At the present time, there exists a vaccination against cholera, consisting of parenteral administration of killed bacteria. Although some countries insist on vaccination against the disease, there are serious doubts as to its real usefulness, given that the current cellular vaccine protects against the consequences of the infection in only 50% of the cases and that the protection is also extremely limited in duration, to less than 6 months.

In Bangladesh, an experimental trial is in progress (1990-92) of an oral vaccine consisting of killed bacteria with the addition of subunit B of cholera toxin, which is known to be highly immunogenic. This product succeeds in inducing lasting protection, without special side effects (Holmgren J., Clemens J., Sack DA., Sanchez J. and Svennerholm AM; "Oral Immunization against cholera" Curr. Top. Microbiol. Immunol. (1988), 146, 197-204).

Cholera toxin resembles the heat labile toxins of enterotoxigenic strains of Escherichia coli in amino acid sequence, structure and mode of action.

The consequences of infection with an enterotoxigenic strain of *E.coli* are similar to, though less serious than, those of cholera, and consist of severe diarrhoea and intestinal disorders.

The CT and LT toxins all comprise a single A subunit (or protomer A) responsible for the enzymic activity of the toxin (herein CT-A or LT-A) and five identical B subunits (or protomer B) which are involved in the binding of the toxin to the intestinal epithelial cells (herein CT-B or LT-B).

The A subunit penetrates the cell membran and causes activation of adenylate cyclase by NAD-dep ndent ADP-ribosylation of a GTP-binding protein which controls the activity of the enzyme. The clinical effect of this is to cause massive fluid loss into the intestine.

Considerable research has been conducted on cholera toxin and the *E. coli* heat labile toxins.

The sequence of CT is known and has been described (Mekalanos J.J. *et al* Nature 306, page 551 (1983)).

The sequence of LT from enterotoxigenic strains of *E.coli* is, as mentioned, 80% homologous to CT and it too is known and described in the scientific literature. Spicer E.K. *et al* (Biol. Chem. 257 p. 5716-5721 (1982)) describe the amino acid sequence of the A sub unit of the heat labile toxin from an enterotoxigenic strain of *E. coli* found in pigs.

A bacterial chromosomal form of LT has been identified and sequenced by Pickett C.L. *et al* (J. Bacteriol. 169, 5180-5187, (1987).

The sequence of the A subunit of LT from a strain of *E. coli* known to affect humans has also been sequenced (Yamamoto *et al*, J. Biol. Chem., 259, 5037-5044, (1984)).

In view of the potential clinical significance of a vaccine against cholera and enterotoxigenic bacteria there is a continuing and great interest in producing a detoxified toxin capable of immunising against cholera and enterotoxigenic bacteria. The techniques of genetic engineering allow specific mutations to be introduced into the genes encoding the toxins and the production of the mutated toxins using now conventional techniques of gene expression and protein purification.

Various groups have attempted to identify mutations of the genes, which involve loss of the toxicity characteristics of the encoded proteins. The studies are predominantly being carried out in respect of the gene for the toxin LT, from *E. coli*.

Harford, S. *et al* (Eur. J. Biochem. 183, page 311 (1989)) describe the production of a toxoid by *in vitro* mutagenesis of the LT-A gene from E.coli pathogenic for pigs. The resulting successful mutation contained a Ser-61-Phe substitution and a Gly-79-Lys substitution, the former being considered the more important. Harford *et al* suggest that, because of the similarities between the LT-A genes in *E.coli* pathogenic to humans and pigs and the CT-A gene, and because the toxins are thought to operate by a common mechanism, it may be possible to produce a cholera holotoxoid by introducing the Ser-61-Phe mutation into the CT-A gene.

Tsuji, T. *et al* (J. Biol. Chem. 265, p. 22520 (1990)) describe the mutation of the LT-A gene from plasmid EWD299 to produce a single substitution Glu-112-Lys which affects the toxicity of the mutant LT yet does not change the immunogenicity of the protein.

Grant, C.C.R. *et al* (Abstract B289 of the 92nd General Meeting of the American Society for Microbiology, 26-30th May 1992) describe conservative substitutions of histidines at 44 and 70 and tryptophan at 127 in LT-A which result in significant reductions in enzymic activity.

Some work has been conducted on mutations to CT.

Kaslow, H.R. *et al* (Abstract B291 of the 92nd General Meeting of the American Society for Microbiology, 26-30th May 1992) describe mutating Asp-9 and His-44 and truncating after amino acid 180 in CT-A which all essentially eliminate activity. Mutating Arg-9 is said to markedly attenuate activity. Mutating other amino acid sites had little effect

on toxicity.

Burnette, W.N. *t al* (Inf. and Immun. 59(11), 4266-4270, (1991)) describe site-specific mutagen sis of CT-A to produce an Arg-7-Lys mutation paralleling that of a known detoxifying mutation in the A subunit of the *Bordetella pertussis* toxin. The mutation resulted in the complete abolition of detectable ADP-ribosyltransferase activity.

International patent application WO 92/19265 (Burnette, Kaslow and Amgen Inc.) describes mutations of CT-A at Arg-7, Asp-9, Arg-11, His-44, His-70 and Glu-112.

Mutations at Glu-110 (LT and CT) and Arg-146 (LT) have also been described in the literature (Lobet, Inf. Immun., 2870, 1991; Lai, Biochem. Biophys. Res. Comm. 341 1983; Okamoto J. Bacteriol. 2208, 1988).

The crystal structure of LT has been determined by Sixma et *al* (Nature, 351, 371-377, May 1991) and confirms the mutatagenesis results described earlier in the literature, explaining structurally the significance of Glu-112 and Ser-61 in activity of the A sub unit and suggesting that His-44, Ser-114 and Arg-54 which are in the immediate neighbourhood may be important for catalysis or recognition.

Summary of the invention

It has now been discovered by further and more detailed analysis of the structure of the toxins that certain further amino acids in the sequences of CT-A and LT-A are in positions capable of decreasing the enzymatic activity of CT and LT when mutated suitably, individually or in conjunction with other mutations.

The object of the present invention is to provide a vaccine which gives total protection against cholera or enterotoxigenic *E. coli*, by m ans of a second generation

product consisting of a single antigen, a toxoid derived from CT or LT, which has been detoxified genetically.

The genetic detoxification of CT or LT retains the immunogenic properties of the toxoid whilst providing a significantly reduced and preferably absent toxicity.

According to a first aspect of the invention there is provided an immunogenic detoxified protein comprising the amino acid sequence of subunit A of a cholera toxin (CT-A) or a fragment thereof or subunit A of an *Escherichia coli* heat labile toxin (LT-A) or a fragment thereof, wherein one or more amino acids at, or in positions corresponding to Val-53, Ser-63, Val-97, Tyr-104 or Pro-106 are replaced with another amino acid.

The replaced amino acids are at locations in the sequences of CT-A or an LT-A which are conserved both in the amino acid sequence and structurally and are thus common to CT and the various LTs.

The immunogenic detoxified protein of the invention adopts substantially the same structural conformation as the wild type naturally occuring toxins. It is immunologically active and cross reacts with antibodies to the wild type toxins.

In this specification, references to CT and LT encompass the various naturally occurring strain variants as well other variants encompassing changes from the sequences disclosed herein which do not affect the immunogenicity of the assembled toxoid.

In this specification, references to amino acid coordinates such as "Val-97" connote the amino acid at that position in the sequence of the mature cholera toxin subunit A (CT-A), that is without the signal sequence (see Figure 1).

Where the specification refers to an LT-A, the amino acid coordinates refer to the corresponding position in CT-A as shown in Figure 1.

Thus, for example, Val-53 in CT corresponds to Val-52 in the LT1 subunit and Ser-63 in CT corresponds to Ser-62 in LT1, there being a single amino acid difference in numbering up to amino acid 89 of the LT1 sequence. Val-97 in the CT sequence corresponds to Val-93 in the LT1 sequence because of the four amino acid difference at that point in the sequence.

In addition, the immunogenic detoxified protein of the invention may include other mutations such as, for example, substitutions at one or more of Arg-7, Asp-9, Arg-11, His-44, Arg-54, Ser-61, His-70, His-107, Glu-110, Glu-112, Ser-114, Trp-127, Arg-146 or Arg-192.

The amino acid substituted for the wild type amino acid may be a naturally occurring amino acid or may a modified or synthetic amino acid. The substitution may involve deletion of an amino acid altogether provided that the mutant retains the necessary immunogenic properties and exhibits a substantially reduced toxicity.

Substitutions which alter the amphotericity and hydrophilicity whilst retaining the steric effect of the substituting amino acid as far as possible are generally preferred.

Preferred substitutions include: Val-53-Asp, Val-53-Glu, Val-53-Tyr, Ser-63-Lys, Val-97-Lys, Val-97-Tyr, His-107-Glu, Tyr-104-Lys, Tyr-104-Asp, Tyr-104-Ser, Pro-106-Ser, Ser-114-Glu, Ser-114-Lys.

As used herein, the term "detoxified" means that the immunogenic composition exhibits a substantially lower toxicity relative to its naturally occurring toxin

counterpart. The substantially lower toxicity should be sufficiently low for the protein to be used in an immunogenic composition in an immunologically effective amount as a vaccine with causing significant side effects. For example, the immunogenic detoxified protein should have a toxicity of less than 0.01% of the naturally occurring toxin counterpart. The toxicity may be measured in mouse CHO cells or preferably by evaluation of the morphological changes induced in Y1 cells. The term "toxoid" means a genetically detoxified toxin.

The immunogenic protein may be a CT or LT subunit A toxoid, but is preferably an assembled toxin molecule comprising a mutated CT-A or LT-A subunit and five B subunits of CT or LT. The B subunit may be a naturally occurring subunit or may itself be mutated.

The immunogenic protein is preferably a naturally occurring CT-A or an LT-A suitably modified as described above. However, conservative amino acid changes may be made which do not affect the immunogenicity or the toxicity of immunogenic protein and preferably do not affect the ability of the immunogenic protein to form complete toxin with B subunit protein. Also, the immunogenic protein may be a fragment of CT-A or an LT-A provided that the fragment is immunogenic and non toxic and contains at least one of the conserved regions containing one of the mutations according to the invention.

According to a second aspect of the invention, there is provided an immunogenic composition for use as a vaccine comprising an immunogenic detoxified protein of the first aspect of the invention and a pharmaceutically acceptable carrier.

The immunogenic composition may additionally contain one or more adjuvants and/or pharmaceutically acceptable diluents.

The invention also provides a vaccine composition comprising an immunogenic detoxified protein according to the first aspect of the invention and a pharmaceutically acceptable carrier. The vaccine composition may further comprise an adjuvant.

According to a third aspect of the invention, there is provided a method of vaccinating a mammal against *Vibrio cholerae* or an enterotoxigenic strain of *Escherichia coli* comprising administering an immunologically effective amount of an immunogenic detoxified protein according to the first aspect of the invention.

The immunogenic detoxified proteins of the invention may be synthesised chemically using conventional peptide synthesis techniques, but are preferably produced by recombinant DNA means.

According to a fourth aspect of the invention there is provided a DNA sequence encoding an immunogenic detoxified protein according to the first aspect of the invention.

Preferably the DNA sequence contains a DNA sequence encoding a complete CT or LT comprising DNA encoding both the detoxified subunit A and subunit B in a polycistronic unit. Alternatively, the DNA may encode only the detoxified subunit A.

According to a fifth aspect of the invention, there is provided a vector carrying a DNA according to the fourth aspect of the invention.

According to a sixth aspect of the invention, there is provided a host cell line transformed with the vector according to the fifth aspect of the invention.

The host cell may be any host capable of producing CT or LT but is preferably a bacterium, most suitably *E.coli* or

V.cholerae suitable engineered to produce the desired immunogenic detoxified protein.

In a further embodiment of the sixth aspect of the invention, the host cell may itself provide a protective species, for example an *E.coli* or *V.cholerae* strain mutated to a phenotype lacking wild type LT or CT and carrying and expressing an immunogenic detoxified protein of the first aspect of the invention.

In a further embodiment of the sixth aspect of the invention the host cell is capable of expressing a chromosomal LT-A gene according to the first aspect of the invention.

According to a seventh aspect of the invention, there is provided a process for the production of an immunogenic detoxified protein according to the first aspect of the invention comprising culturing a host cell according to the sixth aspect of the invention.

According to a eighth aspect of the invention there is provided a process for the production of DNA according to the fourth aspect of the invention comprising the steps of subjecting a DNA encoding a CT-A or an LT-A or a fragment thereof to site-directed mutagenesis.

According to a ninth aspect of the invention there is provided a process for the formulation of a vaccine comprising bringing an immunogenic detoxified protein according to the first aspect of the invention into association with a pharmaceutically acceptable carrier and optionally with an adjuvant.

Industrial Applicability

The immunogenic detoxified protein of the invention constitutes the active component of a vaccine composition useful for the prevention and treatment of cholera

infections or infections by enterotoxigenic strains f *E.coli*. The compositi ns ar thus applicable for us in th pharmaceutical industry.

Brief Description of the Drawings

Figure 1 shows the amino acid sequences of the wild type subunit A from:

i) cholera toxin (CT - Mekalanos *et al* op cit),
ii) heat labile toxin from an *E.coli* strain found in man (LT1_1 - Yamamoto *et al* op cit)
iii) heat labile toxin from an *E.coli* strain found in pigs (LT1- Spicer *et al* op cit), and
iv) heat labile toxin from a chromosomal source (LT1_1 - Pickett *et al* op cit)

The signal sequences are not shown.

In Figure 1, the conventional single letter amino acid code is used. The symbol "." denotes an absent amino acid and acts as a typographical spacer to ensure that the sequences remain in alignment for ease of comparison. The symbol "-" indicates an amino acid in the sequences of LT1 and LT2 which is identical to the corresponding amino acid in CT. The numbers against each line are the amino acid number of the first amino acid on that line.

In Figure 1 the positions of the mutations of the present invention are shown underlined.

Figures 2a and 2b are comparisons of the amino acid and DNA sequences of the A sub units of LT1 and CT.

Figure 3 is a restriction map of plasmid EWD299 (Dallas *et al*), bearing the LT-A gene.

Detailed Description of Embodiments of the Invention

The practice of th pr sent invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Sambrook, et al., MOLECULAR CLONING; A LABORATORY MANUAL, SECOND EDITION (1989); DNA CLONING, VOLUMES I AND II (D.N Glover ed. 1985); OLIGONUCLEOTIDE SYNTHESIS (M.J. Gait ed, 1984); NUCLEIC ACID HYBRIDIZATION (B.D. Hames & S.J. Higgins eds. 1984); TRANSCRIPTION AND TRANSLATION (B.D. Hames & S.J. Higgins eds. 1984); ANIMAL CELL CULTURE (R.I. Freshney ed. 1986); IMMOBILIZED CELLS AND ENZYMES (IRL Press, 1986); B. Perbal, A PRACTICAL GUIDE TO MOLECULAR CLONING (1984); the series, METHODS IN ENZYMOLOGY (Academic Press, Inc.); GENE TRANSFER VECTORS FOR MAMMALIAN CELLS (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory), Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively), Mayer and Walker, eds. (1987), IMMUNOCHEMICAL METHODS IN CELL AND MOLECULAR BIOLOGY (Academic Press, London), Scopes, (1987), PROTEIN PURIFICATION: PRINCIPLES AND PRACTICE, Second Edition (Springer-Verlag, N.Y.), and HANDBOOK OF EXPERIMENTAL IMMUNOLOGY, VOLUMES I-IV (D.M. Weir and C. C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification. All publications ,patents, and patent applications cited herein are incorporated by reference.

In particular, the following amino acid abbreviations are used:

Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic Acid	D	Asp

Cysteine	C	Cys
Glycine	G	Gly
Glutamic Acid	E	Glu
Glutamine	Q	Gln
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

As mentioned above examples of the immunogenic detoxified protein that can be used in the present invention include polypeptides with minor amino acid variations from the natural amino acid sequence of the protein other than at the sites of mutation specifically mentioned.

A significant advantage of producing the immunogenic detoxified protein by recombinant DNA techniques rather than by isolating and purifying a protein from natural sources is that equivalent quantities of the protein can be produced by using less starting material than would be required for isolating the protein from a natural source. Producing the protein by recombinant techniques also permits the protein to be isolated in the absence of some molecules normally present in cells. Indeed, protein compositions entirely free of any trace of human protein contaminants can readily be produced because the only human protein produced by the recombinant non-human host is the recombinant protein at issue. Potential viral agents from natural sources and viral components pathogenic to humans are also avoided. Also, genetically d toxified toxin ar l ss likely to revert

to a toxic from than nor traditional, chemically detoxified toxins.

Pharmaceutically acceptable carriers include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes) and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents (adjuvants).

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: aluminum salts (alum) such as aluminium hydroxide, aluminium phosphate, aluminium sulfate etc., oil emulsion formulations, with or without other specific immunostimulating agents such as muramyl peptides or bacterial cell wall components, such as for example (1) MF59 (Published International patent application WO-A-90/14837, containing 5% Squalene, 0.5% Tween® 80, 0.5% Span® 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA 02164), (2) SAF, containing 10% squalene, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (3) RIBI™ adjuvant system (RAS) (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween® 80 and one or more bacterial cell wall components from the group consisting of monophosphoryl lipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS) preferably MPL+CWS (Detox™), muramyl peptides such as N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-iso-

glutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE) etc., and cytokines, such as interleukins (IL-1, IL-2 etc) macrophage colony stimulating factor (M-CSF), tumour necrosis factor (TNF) etc. Additionally, saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMS (immunostimulating complexes). Furthermore, Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA) may be used. Alum and MF59 are preferred.

The immunogenic compositions (e.g. the antigen, pharmaceutically acceptable carrier and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

Typically, the immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect as discussed above under pharmaceutically acceptable carriers.

Immunogenic compositions used as vaccines comprise an immunologically effective amount of the antigenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, etc.), the capacity of the individual's immune

system to synthesize antibodies, the degre of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

The immunogenic compositions are conventionally administered parenterally, e.g. by injection either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

The term "recombinant polynucleotide" as used herein intends a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of a polynucleotide with which it is associated in nature, (2) is linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature.

The term "polynucleotide" as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA. It also includes known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, f r exampl proteins

(including for e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), thos containing chelat rs (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide.

A "replicon" is any genetic element, e.g., a plasmid, a chromosome, a virus, a cosmid, etc. that behaves as an autonomous unit of polynucleotide replication within a cell; i.e., capable of replication under its own control. This may include selectable markers.

A "vector" is a replicon in which another polynucleotide segment is attached, so as to bring about the replication and/or expression of the attached segment.

"Control sequence" refers to polynucleotide sequences which are necessary to effect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence; in eukaryotes, generally, such control sequences include promoters and transcription termination sequence. The term "control sequences" is intended to include, at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

"Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under c nditions compatibl with the control

sequences.

An "open reading frame" (ORF) is a region of a polynucleotide sequence which encodes a polypeptide; this region may represent a portion of a coding sequence or a total coding sequence.

A "coding sequence" is a polynucleotide sequence which is translated into a polypeptide, usually via mRNA, when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, cDNA, and recombinant polynucleotide sequences.

"PCR" refers to the technique of polymerase chain reaction as described in Saiki, et al., Nature 324:163 (1986); and Scharf et al., Science (1986) 233:1076-1078; and U.S. 4,683,195; and U.S. 4,683,202.

As used herein, x is "heterologous" with respect to y if x is not naturally associated with y in the identical manner; i.e., x is not associated with y in nature or x is not associated with y in the same manner as is found in nature.

"Homology" refers to the degree of similarity between x and y. The correspondence between the sequence from one form to another can be determined by techniques known in the art. For example, they can be determined by a direct comparison of the sequence information of the polynucleotide. Alternatively, homology can be determined by hybridization of the polynucleotides under conditions which form stable duplexes between homologous regions (for example, those which would be used prior to S_1 digestion), followed by digestion with single-stranded specific nuclease(s), followed by size determination of the digested fragments.

As used herein, the term "polypeptide" refers to a polymer of amino acids and does not refer to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not refer to or exclude post expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like. Included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

A polypeptide or amino acid sequence "derived from" a designated nucleic acid sequence refers to a polypeptide having an amino acid sequence identical to that of a polypeptide encoded in the sequence, or a portion thereof wherein the portion consists of at least 3-5 amino acids, and more preferably at least 8-10 amino acids, and even more preferably at least 11-15 amino acids, or which is immunologically identifiable with a polypeptide encoded in the sequence. This terminology also includes a polypeptide expressed from a designated nucleic acid sequence.

The protein may be used for producing antibodies, either monoclonal or polyclonal, specific to the protein. The methods for producing these antibodies are known in the art.

"Recombinant host cells", "host cells," "cells," "cell cultures," and other such terms denote, for example, microorganisms, insect cells, and mammalian cells, that can be, or have been, used as recipients for recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transformed. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due

to natural, accidental, or deliberate mutation. Examples for mammalian host cells include Chines hamster ovary (CHO) and monkey kidney (COS) cells.

Specifically, as used herein, "cell line," refers to a population of cells capable of continuous or prolonged growth and division in vitro. Often, cell lines are clonal populations derived from a single progenitor cell. It is further known in the art that spontaneous or induced changes can occur in karyotype during storage or transfer of such clonal populations. Therefore, cells derived from the cell line referred to may not be precisely identical to the ancestral cells or cultures, and the cell line referred to includes such variants. The term "cell lines" also includes immortalized cells. Preferably, cell lines include nonhybrid cell lines or hybridomas to only two cell types.

As used herein, the term "microorganism" includes prokaryotic and eukaryotic microbial species such as bacteria and fungi, the latter including yeast and filamentous fungi.

"Transformation", as used herein, refers to the insertion of an exogenous polynucleotide into a host cell, irrespective of the method used for the insertion, for example, direct uptake, transduction, f-mating or electroporation. The exogenous polynucleotide may be maintained as a non-integrated vector, for example, a plasmid, or alternatively, may be integrated into the host genome.

By "genomic" is meant a collection or library of DNA molecules which are derived from restriction fragments that have been cloned in vectors. This may include all or part of the genetic material of an organism.

By "cDNA" is meant a complementary DNA sequence that hybridizes to a complementary strand of DNA.

By "purified" and "isolated" is meant, when referring to a polypeptide r nucleotide sequence, that the indicated molecule is present in the substantial absenc of other biological macromolecules of the same type. The term "purified" as used herein preferably means at least 75% by weight, more preferably at least 85% by weight, more preferably still at least 95% by weight, and most preferably at least 98% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000, can be present).

Once the appropriate coding sequence is isolated, it can be expressed in a variety of different expression systems; for example those used with mammalian cells, baculoviruses, bacteria, and yeast.

i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation [Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells." In Molecular Cloning: A Laboratory Manual, 2nd ed.].

Mammalian viral genes are ften highly xpressed and have a

broad host rang ; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples includ the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallotheionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible), depending on the promoter can be induced with glucocorticoid in hormone-responsive cells.

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter [Maniatis et al. (1987) Science 236:1237; Alberts et al. (1989) Molecular Biology of the Cell, 2nd ed.]. Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer [Dijkema et al (1985) EMBO J. 4:761] and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus [Gorman et al. (1982b) Proc. Natl. Acad. Sci. 79:6777] and from human cytomegalovirus [Boshart et al. (1985) Cell 41:521]. Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion [Sassone-Corsi and Borelli (1986) Trends Genet. 2:215; Maniatis et al. (1987) Science 236:1237].

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA mol cule, in which cas th first amin acid at the N-

terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus triparite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation [Birnstiel et al. (1985) *Cell 41*:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In *Transcription and splicing* (ed. B.D. Hames and D.M. Glover); Proudfoot (1989) *Trends Biochem. Sci. 14*:105]. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminater/polyadenylation signals include those derived from SV40 [Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In *Molecular Cloning: A Laboratory Manual*].

Some genes may be expressed more efficiently when introns

(also called intervening sequences) are present. Several cDNAs, however, have been efficiently expressed from vectors that lack splicing signals (also called splice donor and acceptor sites) [see e.g., Gothing and Sambrook (1981) Nature 293:620]. Introns are intervening noncoding sequences within a coding sequence that contain splice donor and acceptor sites. They are removed by a process called "splicing," following polyadenylation of the primary transcript [Nevins (1983) Annu. Rev. Biochem. 52:441; Green (1986) Annu. Rev. Genet. 20:671; Padgett et al. (1986) Annu. Rev. Biochem. 55:1119; Krainer and Maniatis (1988) "RNA splicing." In Transcription and splicing (ed. B.D. Hames and D.M. Glover)].

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 [Gluzman (1981) Cell 23:175] or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a procaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 [Kaufman et al. (1989) Mol. Cell. Biol. 9:946 and pHEBO [Shimizu et al. (1986) Mol. Cell. Biol. 6:1074].

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

ii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art.

Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral

genome are transfected into an insect host cell where th vector and viral genome are allowed to recombine. Th packaged recombinant virus is xpressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to those skilled in the art and fully described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its owned set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, Virology (1989) 17:31.

The plasmid usually also contains th polyhedrin

polyadenylation signal (Miller et al. (1988) Ann. Rev. Microbiol., 42:177) and a pr caryotic ampicillin-r sistance (amp) gene and origin of replication for s lecti n and propagation in E. coli.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: The Molecular Biology of Baculoviruses (ed. Walter Doerfler); EPO Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak et al., (1988), J. Gen. Virol. 69:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) Gene, 73:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the

invertebrate cells and vertebrat c lls, leaders of non-insect origin, such as thos derived from genes encoding human α-interferon, Maeda et al., (1985), Nature 315:592; human gastrin-releasing peptide, Lebacq-Verheyden et al., (1988), Molec. Cell. Biol. 8:3129; human IL-2, Smith et al., (1985) Proc. Nat'l Acad. Sci. USA, 82:8404; mouse IL-3, (Miyajima et al., (1987) Gene 58:273; and human glucocerebrosidase, Martin et al. (1988) DNA, 7:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by in vitro incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb s ction of th baculovirus genom .

Methods for introducing heterologous DNA into the d sired sit in the baculovirus virus ar known in the art. (See Summers and Smith supra; Ju et al. (1987); Smith et al., Mol. Cell. Biol. (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), Bioessays 4:91.The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15 μm in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodi s. "Current Protocols in Microbiology" Vol. 2 (Ausubel

et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, supra; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, inter alia: Aedes aegypti , Autographa californica, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda, and Trichoplusia ni (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) J. Virol. 56:153; Wright (1986) Nature 321:718; Smith et al., (1983) Mol. Cell. Biol. 3:2156; and see generally, Fraser, et al. (1989) In Vitro Cell. Dev. Biol. 25:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith supra.

The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially fr of host debris,

e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

iii. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3") transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps initiate transcription of the lac operon in Escherichia coli (E. coli) [Raibaud et al. (1984) Annu. Rev. Genet. 18:173]. Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (lac) [Chang et al. (1977) Nature 198:1056], and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (trp) [Goeddel et al. (1980) Nuc. Acids Res. 8:4057; Yelverton et al. (1981) Nucl. Acids Res. 9:731; U.S. Patent No. 4,738,921; EPO Publ. Nos. 036 776 and 121 775]. The g-laotamase (bla) promoter system [Weissmann (1981) "The cloning of interferon and other mistakes." In Interferon 3 (ed. I. Gresser)], bacteriophage lambda PL [Shimatake et al. (1981) Nature 292:128] and T5 [U.S. Patent No. 4,689,406] promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter [U.S. Patent No. 4,551,433]. For example, the tac promoter is a hybrid trp-lac promoter comprised of both trp promoter and lac operon sequences that is regulated by the lac repressor [Amann et al. (1983) Gene 25:167; de Boer et al. (1983) Proc. Natl. Acad. Sci. 80:21]. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophase T7 RNA polymerase/promoter system is an example of a coupled promoter system [Studier et al. (1986) J. Mol. Biol. 189:113; Tabor et al. (1985) Proc Natl. Acad. Sci. 82:1074]. In addition, a hybrid promoter can also be comprised of a bacteriophag promoter and an E. coli

operator regi n (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In E. coli, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon [Shine et al. (1975) Nature 254:34]. The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' and of E. coli 16S rRNA [Steitz et al. (1979) "Genetic signals and nucleotide sequences in messenger RNA." In Biological Regulation and Development: Gene Expression (ed. R.P. Goldberger)]. To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site [Sambrook et al. (1989) "Expression of cloned genes in Escherichia coli." In Molecular Cloning: A Laboratory Manual].

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by in vitro incubation with cyanogen bromide or by either in vivo on in vitro incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site f r a processing enzyme (factor

Xa) to cleave the bacteriophage protein from the foreign gene [Nagai et al. (1984) Nature 309:810]. Fusion proteins can also be made with sequences from the lacZ [Jia et al. (1987) Gene 60:197], trpE [Allen et al. (1987) J. Biotechnol. 5:93; Makoff et al. (1989) J. Gen. Microbiol. 135:11], and Chey [EPO Publ. No. 324 647] genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated [Miller et al. (1989) Bio/Technology 7:698].

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria [U.S. Patent No. 4,336,336]. The signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either in vivo or in vitro encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the E. coli outer membrane protein gene (ompA) [Masui et al. (1983), in: Experimental Manipulation of Gene Expression; Ghrayeb et al. (1984) EMBO J. 3:2437] and the E. coli alkaline phosphatase signal sequence (phoA) [Oka et al. (1985) Proc. Natl. Acad. Sci. 82:7212]. As an additional example, the signal sequence of the alpha-amylase gene from various Bacillus

strains can be used to se ret heterologous proteins from B. subtilis [Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EPO Publ. No. 244 042].

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the trp gene in E. coli as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a procaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into th bacterial g nome with an integrating v ctor.

Integrating vectors usually contain at least one sequence homologous to th bacterial chromosome that allows the vector t integrate. Integrations appear to r sult from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various Bacillus strains integrate into the Bacillus chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline [Davies et al. (1978) Annu. Rev.Microbiol. 32:469]. Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, inter alia, the following bacteria: Bacillus subtilis [Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541], Escherichia coli [Shimatake et al. (1981) Nature 292:128; Amann et al. (1985) Gene 40:183; Studier et al. (1986) J. Mol. Biol. 189:113; EPO Publ. Nos. 036 776, 136 829 and 136 907], Streptococcus cremoris [Powell et al. (1988) Appl.

Environ. Microbiol. 54:655]; Streptoc ccus lividans [Powell et al. (1988) Appl. Environ. Microbiol. 54:655], Streptomyces lividans [U.S. Patent No. 4,745,056].

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with $CaCl_2$ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by electroporation. Transformation procedures usually vary with the bacterial species to be transformed. See e.g., [Masson et al. (1989) FEMS Microbiol. Lett. 60:273; Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541, Bacillus], [Miller et al. (1988) Proc. Natl. Acad. Sci. 85:856; Wang et al. (1990) J. Bacteriol. 172:949, Campylobacter], [Cohen et al. (1973) Proc. Natl. Acad. Sci. 69:2110; Dower et al. (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of Escherichia coli with ColE1-derived plasmids. In Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H.W. Boyer and S. Nicosia); Mandel et al. (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; Escherichia], [Chassy et al. (1987) FEMS Microbiol. Lett. 44:173 Lactobacillus]; [Fiedler et al. (1988) Anal. Biochem 170:38, Pseudomonas]; [Augustin et al. (1990) FEMS Microbiol. Lett. 66:203, Staphylococcus], [Barany et al. (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: Streptococcal Genetics (ed. J. Ferretti and R. Curtiss III); Perry et al. (1981) Infec. Immun. 32:1295; Powell et al. (1988) Appl. Environ. Microbiol. 54:655; Somkuti et al. (1987) Proc. 4th Evr. Cong. Biotechnology 1:412, Streptococcus].

iv. Yeast Expression

Yeast expression systems ar als known t ne f ordinary

skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences. Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast PHO5 gene, encoding acid phosphatase, also provides useful promoter sequences [Myanohara et al. (1983) Proc. Natl. Acad. Sci. USA 80:1].

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory

sequences of either the ADH2, GAL4, GAL10, OR PHO5 genes, combined with th transcriptional activation regi n of a glyc lytic enzyme g ne such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, inter alia, [Cohen et al. (1980) Proc. Natl. Acad. Sci. USA 77:1078; Henikoff et al. (1981) Nature 283:835; Hollenberg et al. (1981) Curr. Topics Microbiol. Immunol. 96:119; Hollenberg et al. (1979) "The Expression of Bacterial Antibiotic Resistance Genes i the Yeast Saccharomyces cerevisiae," in: Plasmids of Medical, Environmental and Commercial Importance (eds. K>N> Timmis and A. Puhler); Mercerau-Puigalon et al. (1980) Gene 11:163; Panthier et al. (1980) Curr. Genet. 2:109;].

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by in vitro incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196 056. Another example is a ubiquitin fusion protein. Such a fusion protein is made

with the ubiquitin region that preferably r tains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign pr tein. Through this method, therefore, native foreign protein can be isolated (see, e.g., PCT Publ. No. WO 88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either in vivo or in vitro. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62,096,086) and the A-factor gene (U.S. Patent No. 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recogniz d by yeast are r gulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a procaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 [Botstein et al. (1979) Gene 8:17-24], pCl/1 [Brake et al. (1984) Proc. Natl. Acad. Sci USA 81:4642-4646], and YRp17 [Stinchcomb et al. (1982) J. Mol. Biol. 158:157]. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Enter a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake et al., supra.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homol gous sequences

flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome [Orr-Weaver et al. (1983) Methods in Enzymol. 101:228-245]. An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver et al., supra. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced [Rine et al. (1983) Proc. Natl. Acad. Sci. USA 80:6750]. The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as ADE2, HIS4, LEU2, TRP1, and ALG7, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of CUP1 allows yeast to grow in the presence of copper ions [Butt et al. (1987) Microbiol. Rev. 51:351].

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have b en

developed for transformation into many yeasts. For example, expression vectors have been developed for, inter alia, the following yeasts:Candida albicans [Kurtz, et al. (1986) Mol. Cell. Biol. 6:142], Candida maltose [Kunze, et al. (1985) J. Basic Microbiol. 25:141]. Hansenula polymorpha [Gleeson, et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302], Kluyveromyces fragilis [Das, et al. (1984) J. Bacteriol. 158:1165], Kluyveromyces lactis [De Louvencourt et al. (1983) J. Bacteriol. 154:737; Van den Berg et al. (1990) Bio/Technology 8:135], Pichia guillerimondii [Kunze et al. (1985) J. Basic Microbiol. 25:141], Pichia pastoris [Cregg, et al. (1985) Mol. Cell. Biol. 5:3376; U.S. Patent Nos. 4,837,148 and 4,929,555], Saccharomyces cerevisiae [Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1929; Ito et al. (1983) J. Bacteriol. 153:163], Schizosaccharomyces pombe [Beach and Nurse (1981) Nature 300:706], and Yarrowia lipolytica [Davidow, et al. (1985) Curr. Genet. 10:380471 Gaillardin, et al. (1985) Curr. Genet. 10:49].

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz et al. (1986) Mol. Cell. Biol. 6:142; Kunze et al. (1985) J. Basic Microbiol. 25:141; Candida]; [Gleeson et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302; Hansenula]; [Das et al. (1984) J. Bacteriol. 158:1165; De Louvencourt et al. (1983) J. Bacteriol. 154:1165; Van den Berg et al. (1990) Bio/Technology 8:135; Kluyveromyces]; [Cregg et al. (1985) Mol. Cell. Biol. 5:3376; Kunze et al. (1985) J. Basic Microbiol. 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; Pichia]; [Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1929; Ito et al. (1983) J. Bacteriol. 153:163 Saccharomyces]; [Beach and Nurse (1981) Nature 300:706; Schizosaccharomyces]; [Davidow et al. (1985) Curr. Genet.

10:39; Gaillardin et al. (1985) Curr. Genet. 10:49; Yarrowia].

Example 1 - Detoxified LT

A fragment of the gene for LT was extracted from plasmid EWD299 [Dallas W.S., Gill D.M. and Falkow S., 1979, J. Bacteriol., 139, 850-858] by digestion with the restriction enzymes SmaI and EcoRI, and was recloned in the vector Bluescript KS suitable for producing single strands of DNA [Sambrook J., Fritsch E. and Maniatis, T. "Molecular Cloning", Cold Spring Harbor].

BW313 cells were transformed by the clones thus obtained and allowed to grow for 14 hours in a culture medium consisting of Luria Broth with the addition of 1 µg/ml of uridine.

A series of synthetic oligonucleotides (listed in Table 1 below), containing the mutation, or the desired bases instead of the natural ones, and a sequence of 10 bases upstream and 10 downstream of the same mutation, identical to the natural ones, was first of all synthesised chemically and then phosphorylated, 1.5 pmol thereof being treated at 37°C with 5 units of kinase.

After halting the reaction with a 100 mM EDTA solution, the oligonucleotides were annealed to the single strand containing the LT gene, by heating for 5 minutes at 70°C and cooling slowly for about one hour in ice.

At that stage there was added to this cold solution (25 µl) a solution of free nucleotides, the enzyme DNA ligase and the enzyme DNA polymerase, in a final volume of 100 µl.

The solution thus obtained was kept for five minutes in ice, five minutes at ambient temperature and two hours at 37°C.

Suitable cells of E. coli were transformed with the reacti n

mixture, in accordance with the usual techniques [Sambrook J., Fritsch E. and Maniatis T. "Molecular Cloning" Cold Spring Harbor], and the site-directed mutagenesis was checked by sequencing of the clones obtained.

The SmaI-EcoRI fragment containing the various mutations was substituted for the original SmaI-EcoRI insert in the plasmid EWD299.

The strains which encode the mutated toxin were then grown in 10 ml of Luria Broth for 12 hours at 37°C.

The cultures were centrifuged and the precipitate containing the cells was resuspended in 300 ml of a solution containing 25% of sucrose and 50 mM of Tris buffer at pH8, and the mixture was treated for one hour at ambient temperature with 1 mg/ml of a solution of Polymixin B.

The presence of the toxoid in the periplasmatic supernatant liquor was verified by means of Western Blot and its toxicity was evaluated by the inducement or lack of inducement of morphological changes in Y1 cells (see Table 1).

Y1 cells are adrenal tumour epithelial cells which become markedly more rounded when treated with a solution containing CT or LT [Yasamure Y., Buonassisi V. and Sato G., "Clonal analysis of differentiated function in animal cell cultures", Cancer Res., 1966, 26, 529-535]. The toxicity of CT and LT is correlated with this morphological transition. The periplasmic supernatant is diluted with a solution of F10 medium, horse serum 1.5%, glutamine and gentamycin to lesser and lesser concentrations and Y1 cells (250000 cells/ml) are incubated with the resulting solutions for 48 hours at 37°C under an atmosphere of CO_2. The morphology of the cells is evaluated.

In all cases, immunogenicity was shown by correct assembly

of the complete toxoid and by cross reaction of the toxoid with antibody to the wild type LT.

The results are shown in Table I below.

In this Table (and in Table II below) the toxicity symbols mean as follows:

+++ toxic after dilution 1:2000 (wild type toxicity)

++ toxic up to dilution 1:250

+ toxic up to dilution 1:64

- not toxic, even undiluted

TABLE I

Example	Mutation	Oligonucleotide Sequence	Toxicity
1.1 LT	Val-53-Asp	291-ACCGGCTTTGATAGATATGAT-311	-
1.2 LT	Val-53-Glu	291-ACCGGCTTTGAAAGATATGAT-311	-
1.3 LT	Val-53-Tyr	291-ACCGGCTTTTACAGATATGAT-311	-
1.4 LT	Ser-63-Lys	322-GTTTCCACTAAGCTTAGTTTG-342	-
1.5 LT	Val-97-Lys	424-ATGTTTAATAAGAATGATGTA-444	-
1.6 LT	Val-97-Tyr	424-ATGTTTAATTACAATGATGTA-444	-
1.8 LT	His-107-Glu	454-TACAGCCCTGAGCCATATGAA-474	+/-
1.9 LT	Tyr-104-Lys	445-ATTAGCGTAAAGAGCCCT-462	-
1.10 LT	Tyr-104-Asp	445-ATTAGCGTAGATAGCCCT-462	-
1.11 LT	Tyr-104-Ser	447-TAGCGTAAGTAGCCCTCA-464	-
1.12 LT	Pro-106-Ser	453-ATACAGCAGCCACCCATA-470	-

Two mutation of serine (Ser-114-Glu:477-GGAGGTGAAGCGTTAGG-494 and Ser-114-Lys:477-GGAGGTTAAAGCGTTAGG-494) were also shown to exhibit substantially reduced toxicity.

Comparative Examples

A LT	LT Wild Type		+++
B LT	Arg-210-Asp	769-ATATATCTCAACGAATATCAA-789	+
C LT	Leu-41-Phe	113-ATATTAATTTCTATGATC-130	NA
D LT	His-44-Phe	121-CTTTATGATTTTGCGAGA-138	NA
E LT	Ala-45-Tyr	125-ATGATCACTATAGAGGAA-142	NA
F LT	Arg-54-Ala	152-GCTTTGTCGCGTATGATG-169	++
G LT	Arg-54-Lys	151-GGCTTTGTCAAGTATGATGAT-171	++
H LT	Tyr-59-Met	167-ATGACGGAATGGTTTCCA-184	+++
I LT	Val-60-Gly	169-GACGGATATGGATCCACTTCT-189	NA
J LT	Ser-68-Lys	193-AGTTTGAGAAAGGCTCACTTA-213	++
K LT	Ser-68-Pro	193-AGTTTGAGACCAGCTCACTTA-213	NA
L LT	His-70-Pro	199-AGAAGTGCTCCTTTAGCAGGA-219	NA
M LT	Ala-72-Arg	205-GCTCACTTAAGGGGACAGTCT-225	++
N LT	Ala-72-His	205-GCTCACTTACATGGACAGTCT-225	+++
LT	Arg-192-Asn	565-GATTCATCAATTACAATCACA-585	+++

(NA means "not assembled", i. the holotoxin AB_5 is not formed at all)

Example 2 - Detoxified CT

The procedure followed in the case of the gene for the toxin CT is analogous to that described above.

A fragment containing the gene for a CT was amplified by means of the polymerase chain reaction (PCR) technique from plasmid pCT322. An alternative and equivalent source of the CT gene is plasmid pJM17 (Pearson *et al*, PNAS USA, 79, (1982), 2976-2980).

The following two synthetic primers were used:

1) GGCAGATTCTAGACCTCCTGATGAAATAAA

2) TGAAGTTTGGCGAAGCTTCTTAATTTGCCATACTAATTGCGGCAATCGCAT

containing respectively an XbaI site and an artificial HindIII site (shown underlined).

The resulting amplified fragment, XbaI-HindIII, which has a length of 1074 base pairs, contains the codons of the two sub-units, A and B, but not the sequence encoding the leader peptide of the A sub-unit. This fragment was recloned in Bluescript KS vector and was treated in accordance with the procedure described above for LT, so as to effect the site-directed mutagenesis.

TABLE II

Example	Mutation	Oligonucleotide Sequence	Toxicity
2.1	CTVal-53-Asp	ACGGGATTTGACAGGCACGAT	-
2.2	CTSer-63-Lys	GTTTCCACCAAGATTAGTTTG	-
2.3	CTVal-97-Lys	ATGTTTAACAAGAATGATGTA	-
2.4	CTSer-106-Pro	GGCATACAGTAGCCATCCAGA	-
Comparative Examples			
A	CT Arg-192-Asn	GAATGCTCCAAACTCATCGAT	+++
B	CTArg-54-His	GGATTTGTTCATCACGATGAT	++

The following mutations also proved to abolish toxicity: His-107-Asn (TACAGTCCTAACCCAGATGAA), Glu-110-Ser (TCATCCAGATTCGCAAGAAGT), Glu-112-Ala (CAGATGAACAAGCTGTTTCTG) and Ser-114-Glu (CAAGAAGTTGAAGCTTTAGGT).

It will be understood that the invention is described above by way of example only and modifications of detail may be made within the scope and spirit of the invention.

CLAIMS:

1. An immunogenic detoxified protein comprising the amino acid sequence of subunit A of a cholera toxin (CT-A) or a fragment thereof or the amino acid sequence of subunit A of an *Escherichia coli* heat labile toxin (LT-A) or a fragment thereof wherein one or more amino acids at, or in positions corresponding to Val-53, Ser-63, Val-97, Tyr-104 or Pro-106 are replaced with another amino acid.

2. An immunogenic detoxified protein according to claim 1 wherein additionally one or more amino acids at, or in positions corresponding to Arg-7, Asp-9, Arg-11, His-44, Arg-54, Ser-61, His-70, His-107, Glu-110, Glu-112, Ser-114, Trp-127, Arg-146 or Arg-192 are replaced.

3. An immunogenic detoxified protein according to claim 1 or 2 comprising one or more of the following amino acid replacements Val-53-Asp, Val-53-Glu, Val-53-Tyr, Ser-63-Lys, Val-97-Lys, Val-97-Tyr, His-107-Glu, Tyr-104-Lys, Tyr-104-Asp, Tyr-104-Ser, Pro-106-Ser, Ser-114-Glu, Ser-114-Lys.

4. An immunogenic composition for use a vaccine comprising an immunogenic detoxified protein according to any one of the preceding claims and a pharmaceutically acceptable carrier.

5. A vaccine composition comprising an immunogenic detoxified protein according to any one of claims 1 to 3 and a pharmaceutically acceptable carrier.

6. A vaccine composition according to claim 5 further comprising an adjuvant.

7. A DNA sequence encoding an immunogenic detoxified protein according to any one of claims 1 to 3.

8. A vector carrying a DNA according to claim 7.

9. A host cell line transformed with th vector according to claim 8.

10. A process for the production of an immunogenic detoxified protein according to any one of claims 1 to 3 comprising culturing a host cell according to claim 9.

11. A process for the production of a DNA according to claim 7 comprising the steps of subjecting a DNA encoding a CT-A or an LT-A or a fragment thereof to site-directed mutagenesis.

12. A method of vaccinating a mammal against *Vibrio cholerae* or an enterotoxigenic strain of *Escherichia coli* comprising administering an immunologically effective amount of an immunogenic detoxified protein according to any one of claims 1 to 3.

13. A process for the formulation of a vaccine according to claim 5 comprising bringing an immunogenic detoxified protein according to any one of claims 1 to 3 into association with a pharmaceutically acceptable carrier.

14. A process for the formulation of a vaccine according to claim 6 comprising bringing an immunogenic detoxified protein according to any one of claims 1 to 3 into association with an adjuvant.

```
LT2      1  -.-.FF-----T----R-A---L---QQ-AYE---PI---   38
LT1      1  ----------------FRS-----.----------------   39
LT1_1A   1  -G-R-------------R--------HN--------------   40
CT       1  NDDKLYRADSRPPDEIKQSGGLMPRGQSEYFDRGTQMNIN   40

                --E-----V--NT--N------TVT--Q---I--N--GS-   78
                ------------------------------Y-----      79
                -------------Y--------L-------A--S----Y   80
                LYDHARGTQTGFVRHDDGYVSTSISLRSAHLVGQTILSGH   80

                NE-----V-P---L-D--G---R---Y-S-N-FA------   118
                -LTIYI---...---------IS-----------------   116
                ---------------------V-----Y-----------   120
                STYYIYVIATAPNMFNYNDVLGAYSPHPDEQEVSALGGIP   120

                L---I-----SF-A-EGGMQ---D--GDLF-G-TV--N--   158
                ------------------------------------     156
                ----------N---I--R-----E------R--N----E-   160
                YSQIYGWYRVHFGVLDEQLHRNRGYRDRYYSNLDIAPAAD   160

                --Q-----SNFP----M--STF--EQ-VPNNKEFK-GV-I   198
                ------------------------------------     196
                --R------D-Q-----------Q---DSS-TITGD--N   200
                GYGLAGFPPEHRAWREEPWIHHAPPGCGNAPRSSMSNTCD   200

           SA-NV--KYD-MNFKKLL--RLALTFFM--D-F-GVHGE----   241
           ------------------------.-----.-----.----   236
           -E--N-STIY-R----------D-.--EV-.IY---.R---   240
           EKTQSLGVKFLDEYQSKVKRQIFSGY.QSDID.THNRI.KDEL   240
```

Figure 1

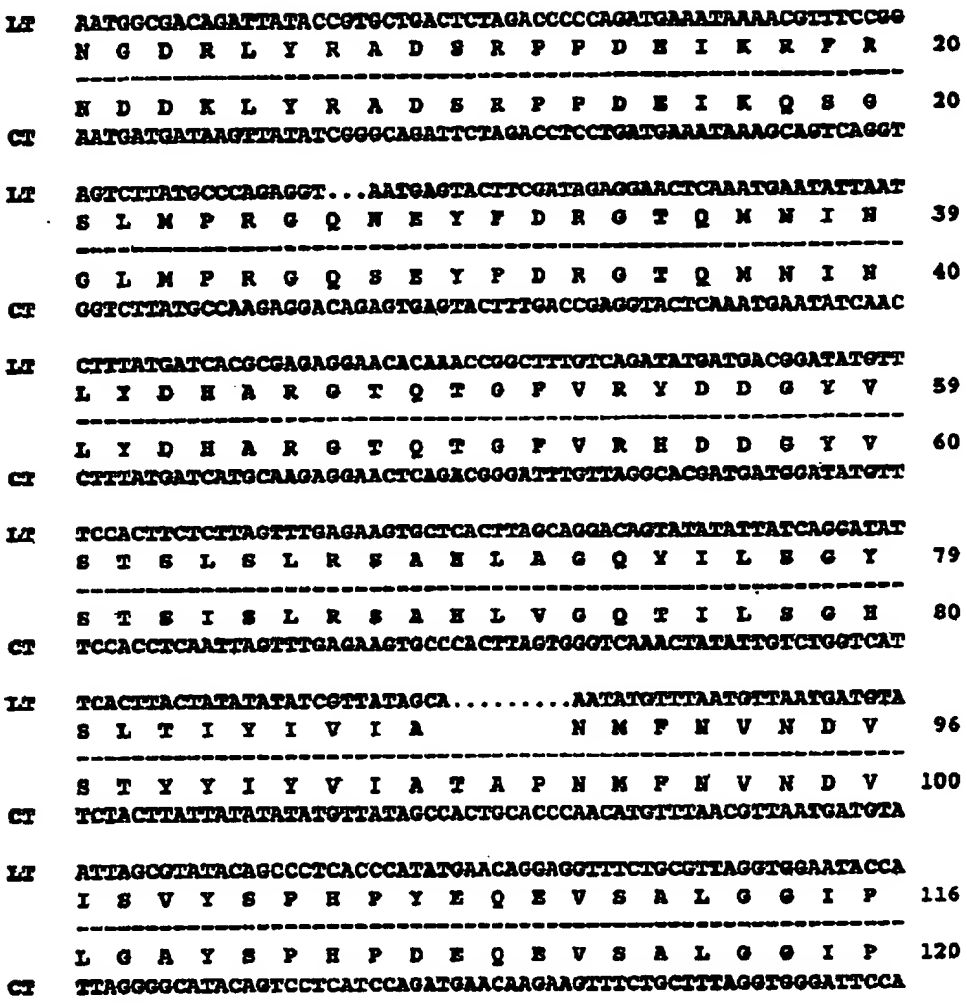

LT AATGGCGACAGATTATACCGTGCTGACTCTAGACCCCCAGATGAAAATAAAACGTTTCCGG
N G D R L Y R A D S R P P D E I K R F R 20
N D D K L Y R A D S R P P D E I K Q S G 20
CT AATGATGATAAGTTATATCGGGCAGATTCTAGACCTCCTGATGAAATAAAGCAGTCAGGT
LT AGTCTTATGCCCAGAGGT...AATGAGTACTTCGATAGAGGAACTCAAATGAATATTAAT
S L M P R G Q N E Y F D R G T Q M N I N 39
G L M P R G Q S E Y P D R G T Q M N I N 40
CT GGTCTTATGCCAAGAGGACAGAGTGAGTACTTTGACCGAGGTACTCAAATGAATATCAAC
LT CTTTATGATCACGCGAGAGGAACACAAACCGGCTTTGTCAGATATGATGACGGATATGTT
L Y D H A R G T Q T G F V R Y D D G Y V 59
L Y D H A R G T Q T G F V R H D D G Y V 60
CT CTTTATGATCATGCAAGAGGAACTCAGACGGGATTTGTTAGGCACGATGATGGATATGTT
LT TCCACTTCTCTTAGTTTGAGAAGTGCTCACTTAGCAGGACAGTATATATTATCAGGATAT
S T S L S L R S A H L A G Q Y I L S G Y 79
S T S I S L R S A H L V G Q T I L S G H 80
CT TCCACCTCAATTAGTTTGAGAAGTGCCCACTTAGTGGGTCAAACTATATTGTCTGGTCAT
LT TCACTTACTATATATATCGTTATAGCA.....AATATGTTTAATGTTAATGATGTA
S L T I Y I V I A N M F N V N D V 96
S T Y Y I Y V I A T A P N M F N V N D V 100
CT TCTACTTATTATATATATGTTATAGCCACTGCACCCAACATGTTTAACGTTAATGATGTA
LT ATTAGCGTATACAGCCCTCACCCATATGAACAGGAGGTTTCTGCGTTAGGTGGAATACCA
I S V Y S P H P Y E Q E V S A L G G I P 116
L G A Y S P H P D E Q E V S A L G G I P 120
CT TTAGGGGCATACAGTCCTCATCCAGATGAACAAGAAGTTTCTGCTTTAGGTGGGATTCCA

Figure 2a

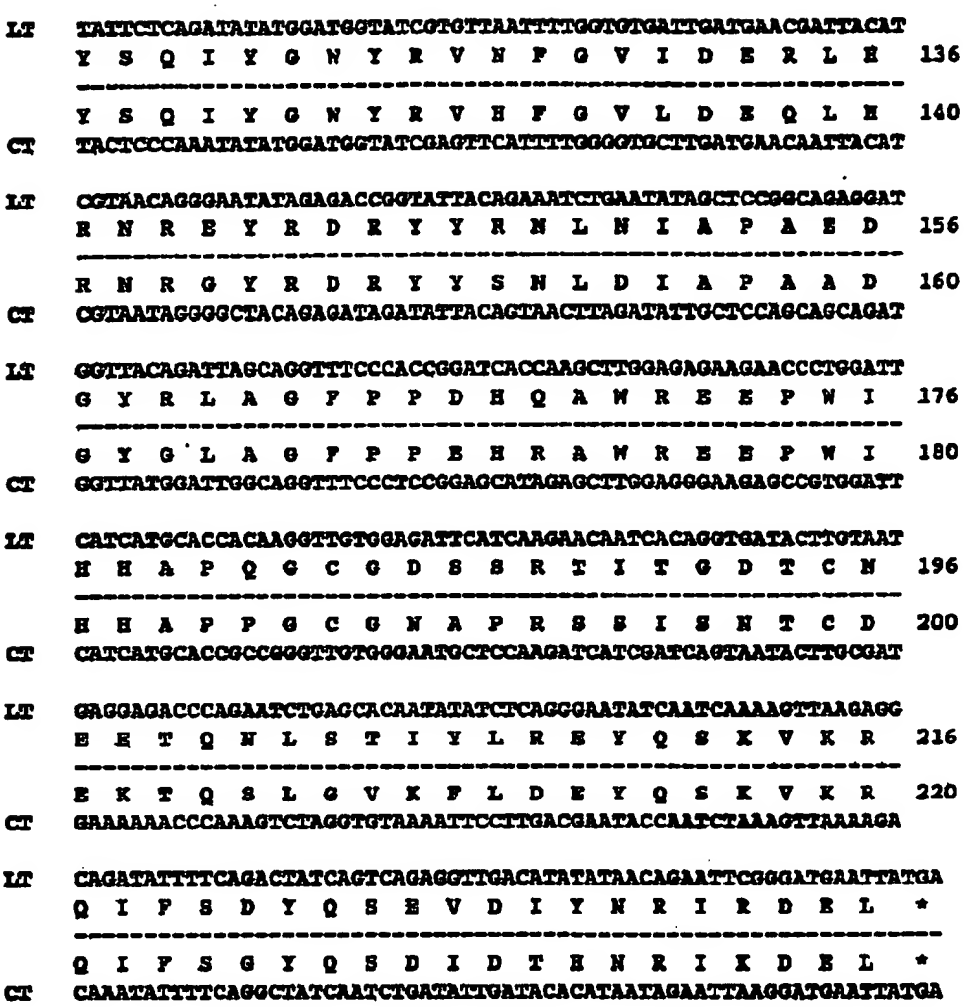


Figure 2b

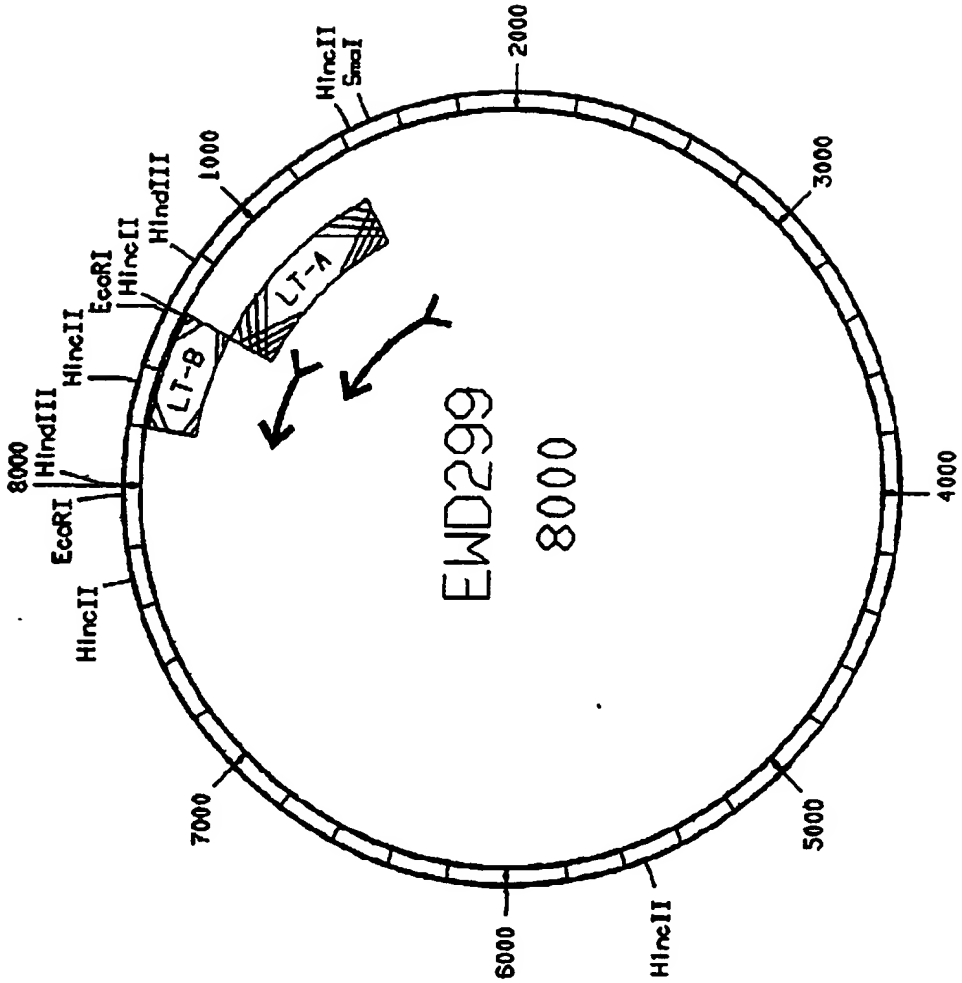

EWD299
8000
LT-A
LT-B
HincII
SmaI
HindIII
EcoRI
HincII
HincII
HindIII
EcoRI
HincII
1000
2000
3000
4000
5000
6000
7000
8000

Figure 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/[redacted]92/03016

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claim 12 is directed to a method of treatment of the human /animal body the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	INFECTION AND IMMUNITY vol. 59, no. 11, November 1991, AM. SOC. MICROBIOL., BALTIMORE, US; pages 4266 - 4270 W.N. BURNETTE ET AL. 'Site-specific mutagenesis of the catalytic subunit of cholera toxin: Substituting lysine for arginine 7 causes loss of activity' cited in the application see page 4266, right column, line 14 - line 29	1,2,7,11
Y	J. BACTERIOLOGY vol. 170, no. 5, May 1988, AM. SOC. MICROBIOL., BALTIMORE, US; pages 2208 - 2211 K. OKAMOTO ET AL. 'Effect of substitution of glycine for arginine at position 146 of the A1 subunit on biological activity of Escherichia coli heat-labile enterotoxin' see page 2208, right column, line 18 - page 2209, left column, line 7	1,2,7,11
P,Y	WO,A,9 219 265 (AMGEN INC.) 12 November 1992 cited in the application see page 22, line 25 - page 27, line 26	1,2,7,11

INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 92/03016

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/00; C12N15/31; A61K39/106; A61K39/108
C12P21/02; C12N1/21; //(C12N1/21,C12R1:19)

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C12N ; C12P ; A61K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category[10]	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
Y	INFECTION AND IMMUNITY vol. 59, no. 9, September 1991, AM. SOC. MICROBIOL.,BALTIMORE, US; pages 2870 - 2879 Y. LOBET ET AL. 'Effect of site-directed mutagenic alterations on ADP-ribosyltransferase activity of the a subunit of Escherichia coli heat-labile enterotoxin' cited in the application see page 2871, right column, line 7 - page 2873, right column, line 5 -/---	1,2,7,11

* Special categories of cited documents: [10]
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
20 APRIL 1993	10. 05. 93
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	HORNIG H.

Form PCT/ISA/210 (second sheet) (January 1985)

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.

EP 9203016
SA 69235

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 20/04/93

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9219265	12-11-92	AU-A- 2293092	21-12-92